DOCUMENT RESUME

ED 053 130 TE 002 519

AUTHOR Williams, Frederick; Whitehead, Jack L.
TITLE Language In the Classroom: Studies of the Pygmalion Effect.
PUB DATE Apr 71
NOTE 6p.
JOURNAL CIT English Record; v21 n4 p108-113 Apr 1971

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS Data Collection, Educational Research, *Nonstandard Dialects, *Rating Scales, Socioeconomic Status, *Tape Recordings, *Teacher Attitudes

ABSTRACT

The present research involves teachers' attitudes. In the first phase, individual teachers from inner city schools were interviewed according to a procedure whereby selected samples of tapes of various dialects were played and the teacher was asked to describe her impressions of the child. Adjectives from these free-responses formed the basis for the development of rating scales that were eventually used to obtain quantitiative data on teachers' attitudes. In subsequent phases of the research, sample groups of teachers from the same population were administered selected tapes from the above materials, which they then rated on scales like the one above. Another series of studies used videotapes of childrens' speech samples. The tapes included samples of children from Anglo, Black, and Mexican-American families and within each of these groups, children from low or middle status families. The adjectives used by small samples of teachers who viewed and described their impressions of the children in the videotapes were again used as a basis for developing rating scales. Just as the present studies imply a bias in teachers' attitudes toward nonstandardness, they also suggest ways to measure such bias and still more to gauge the effects of teacher training. (CK)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ED053130

English Record
Vol 21, No. 2,
April, 1971.

LANGUAGE IN THE CLASSROOM: STUDIES OF THE PYMALION EFFECT

Frederick Williams
Jack L. Whitehead

Although Shaw's *Pygmalion*, as well as our everyday intuitions, make us well aware of the relationship between speech characteristics and social attitudes, only in the last decade has this relation been examined as a topic of behavioral sciences research. Perhaps the best known studies along this line have come from the work of the social psychologist Wallace Lambert and his associates at McGill University.[1] Among such studies have been experiments where, for example, listeners would assign personality characteristics thought to be associated with speakers of tape-recorded examples of French and English speech. Unknown to the listeners was that the samples spoken in the two languages were earlier recorded by persons who were perfect and coordinate bilinguals. Thus it was interesting, if not amusing, to find that, for example, listeners rated English speakers as better looking or more intelligent than their French-speaking counterparts, or that the French speakers were more kind or religious.

[1] "White People's Time, Colored People's Time," **The Disadvantaged Learner: Knowing, Understanding, Evaluating,"** Staten W. Webster, ed. (San Francisco, Chandler Publishing Co., 1966), pp. 187-190.

[1] W. E. Lambert, R. C. Hodgson, R. C. Gardner, and S. Fillenbaum, "Evaluational Reactions to Spoken Languages," **Journal of Abnormal and Social Psychology,** LX (January, 1960), 44-51; M. Anisfeld, N. Bogo, and W. E. Lambert, "Evaluational Reactions to Accented English Speech," **Journal of Abnormal and Social Psychology,** LXV (October, 1962), 223-231; E. Anisfeld and W. E. Lambert, "Evaluational Reactions of Bilingual and Monolingual Children to Spoken Language," **Journal of Abnormal and Social Psychology,** LXIX (July, 1964), 89-97; W. E. Lambert, M. Anisfeld, and Grace Yeni-Komshian, "Evaluational Reactions of Jewish and Arab Adolescents to Dialect and Language Variations," **Journal of Personality and Social Psychology,** II (July, 1965), 84-90; and W. E. Lambert, Hannah Frankel, and G. R. Tucker, "Judging Personality Through Speech: A French-Canadian Example," **Journal of Communication,** XVI (December, 1966), 305-321.

Professor Frederick Williams is Director of the Center for Communication Research in the School of Communication at the University of Texas at Austin. Prior to his coming to Texas, he was a member of the senior research staff of the Institute for Research on Poverty at the University of Wisconsin. Dr. Williams is the author of Reasoning with Statistics, Language and Poverty: Perspectives on a Theme *and the forthcoming* Language and Speech: Basic Perspectives. *Dr. Jack L. Whitehead is a member of the senior research staff of the Center for Communication Research and Assistant Professor of Speech in the School of Communication. Professor Whitehead is the author of several research articles on language and attitude and on small group communication.*

Copyright New York State English Council 1971.

To anyone who has worked with minority group children in the classroom, there is the question of the degree to which the speech of such children elicits social stereotyping in the mind of a teacher, and whether such stereotyping and associated attitudes might mediate the teacher's instructional behavior toward the child. This latter question, of course, relates to the theme of Rosenthal and Jacobson's now well-known *Pygmalion in the Classroom.*[2] These researchers found a reliable correlation between experimentally manipulated attitudes of teachers toward children and the subsequent progress of those children in their classes.

The research to be summarized in this paper reflects the first step in the above relation—that is, the degree to which the speech characteristics of children can be related to the attitudes of teachers.

Background

One main study, the details of which are reported elsewhere,[3] led to the present research involving teachers' attitudes. Sound tapes of speech samples of Black and White, male and female, fifth and sixth grade children sampled from low and middle income families served as stimulus materials in this initial study. These tapes represented variations among degrees of Negro-nonstandard English as well as White children's speech variations relative to standard English as recorded in conversations between a linguistic fieldworker and the child. The conversations were in response to two initial questions and followup inquiries: "What kinds of games do you play around here?" and "What are your favorite television programs?"

In the first phase of this research, individual teachers from inner-city schools in Chicago were interviewed according to a procedure whereby selected samples of the above tapes were played, then the teacher was asked to describe her impressions of the child—e.g., his ethnicity, educational background, language and speech characteristics, and so on. Adjectives from these free-responses formed the basis for the development of rating scales that were eventually used to obtain quantitative data on teachers' attitudes. A sample of one such scale appears as follows:

THE CHILD SEEMS: hesitant—:—:—:—:—:—:—enthusiastic

In subsequent phases of the research, sample groups of teachers from the same population were administered selected tapes from the above materials which they then rated on scales like the one above. The result of this series of projects was that although teachers would use some 22 individual scales in rating children's speech, their ratings were generally symptomatic of only two main evaluative dimensions.[4] One of these dimensions was labeled as *confidence-eagerness,* which was a reflection of highly similar ratings on adjectival scales such as "The child seems: unsure—confident" and ". . . reticent—eager." The second main dimension of evaluation was labeled as *ethnicity-nonstandardness,* as identified from the apparently global ratings given on such scales as "Pronunciation is: standard—nonstandard" or "Grammar is: good—bad," and so on. In sum, the evidence pointed to the generalization that teachers typically gave

[2] Robert Rosenthal and Lenore Jacobson, **Pygmalion in the Classroom** (New York, 1968).

[3] Frederick Williams, "Psychological Correlates of Speech Characteristics: On Sounding 'Disadvantaged,'" **Journal of Speech and Hearing Research,** XIII (September, 1970), 472-488.

[4] For the statistically inclined, these dimensions were obtained by a factor analysis of the scale intercorrelations. Such an analysis asks whether persons' uses of the detailed scales reflect more global judgments as indicated by very similar (correlated) uses of selected scales. If factors can be found, it is the researcher's task to interpret and label them (if possible).

rather global evaluations of children's speech, but that these evaluations were along at least two relatively independent dimensions.

That the above judgmental dimensions had been validly identified was supported by two types of information. First, the children who had been selected from the low and middle income groups, and whose speech reflected this social stratification, were reliably differentiated on the average along the above two dimensions. That is, the middle status children were typically rated as less nonstandard and ethnic-sounding, and as more confident and eager than their lower status counterparts. Second, it was eventually possible to predict mathematically teachers' ratings on the two dimensions by using several characteristics of the speech samples as predictor variables. Thus, for example, deviations from standard English in pronunciation or use of /s/, /z/, /Θ/, or / ð / phonemes and pronominal apposition were salient predictors of nonstandard ratings, and the lack of hesitation phenomena was an effective predictor of confidence-eagerness ratings.

One unexpected finding in this early research was that individual teacher-raters were sometimes quite consistent with themselves in terms of their ratings of all Black children or all White children, apart from the actual income group of the children or even details of their speech samples. The latter point, of course, reflects a lack of predictability of ratings as based upon quantified characteristics of the samples. This phenomenon suggested that some teacher-raters (or some teachers in some ratings) may have been reporting simply stereotyped judgments of a category of child rather than their detailed perceptions and evaluations of what was presented on the stimulus tapes. Although the results are too detailed to be summarized here, further evidence of the stereotype biases of individual raters was revealed in a companion study[5] (using the same data) where it was found that various teacher-raters could be reliably grouped together on the basis of their commonality in rating certain types of children on certain scales.

The consistent and readily interpretable results of this earlier research prompted two main questions for further study:[6]

1. What is the generality of the two-dimensional judgmental model when the teacher can see as well as hear the child-speaker?
2. What is the relation between a teacher's ratings of children and her stereotypes of children of different income groups and ethnicities?

Generality of the Judgmental Model

In a new series of studies, the same technique for deriving rating scales as described above was undertaken, but this time, videotapes rather than simply audio tapes of children's speech samples were employed. These videotapes were obtained by interviewing children from specified income and ethnic neighborhoods of Austin, Texas and its environs. The tapes included samples of children from Anglo, Black, and Mexican-American families,[7] and within each of these groups, children from low or middle-status

---

[5] Rita C. Naremore, "Teachers' Judgments of Children's Speech: A Factor Analytic Study of Attitudes" unpubl. Ph.D. diss. (University of Wisconsin, 1969).

[6] Technical reports of these studies are available from ERIC (numbers not yet assigned) as: Frederick Williams, Jack L. Whitehead, and Jane Traupmann, "Semantic Differential Scaling of Audiovisual Recordings of Children's Speech Samples," technical report, (Austin, 1970a); Frederick Williams, Jack L. Whitehead, and Jane Traupmann, "Latency of Teachers' Semantic Differential Ratings of Children's Speech," technical report, (Austin, 1970b); Frederick Williams, Jack L. Whitehead, and Jane Traupmann, "Correspondence Between Semantic Differential Ratings of Children's Speech and Speech Anticipated upon the Basis of Stereotype," technical report, (Austin, 1970c).

[7] The authors are aware of the problems involved in labeling ethnic groups. Since these labels are used in our technical reports of this research, we have chosen to use them again here.

families. Initially, six children were interviewed for each ethno-status category. The interviews were conducted in a living-room-like atmosphere by a 24 year-old Anglo female fieldworker. Each interview was from five to ten minutes in length and was guided by the fieldworker's questions about games and television (as in the earlier studies).

The adjectives used by small samples of teachers who viewed and described their impressions of the children in the videotapes were again used as a basis for developing rating scales. An experimental set of 59 scales was used by teacher-raters in a testing design whereby ratings were obtained of children in the six different ethno-status categories and each in a video-only, audio-only, and audio-video presentation mode. In this phase of the research, teacher-raters also had the opportunity to omit the use of any individual scales they thought irrelevant to their judgments.

Statistical analyses of all ethno-status and presentation modes combined indicated that the data fit a two-dimensional judgmental model nearly identical to the earlier model. The dimensions and their main correlated scales were as follows:

Confidence-eagerness

The child seems: active—passive (.86) [8]
The child seems to: enjoy talking—dislike talking (.85)
The child seems: hesitant—enthusiastic (.84)
The child seems: shy—talkative (.83)
The child seems: eager to speak—reticent to speak (.83)

Ethnicity-nonstandardness

The language shows: a standard American style—marked ethnic style (.90)
The language spoken in the child's home is probably: standard American style—marked ethnic style (.90)
The child seems culturally: advantaged—disadvantaged (.80)
The child's family is probably: high social status—low social status (.80)
Pronunciation is: standard—nonstandard (.70)

Separate analyses of ratings of children in each of the ethnic categories revealed nearly identical versions of the above results, thus testifying to the generality of the judgmental model across at least the three ethnic groups. Separate analyses of each of the presentation modes also yielded evidence of the two-factor model. However, the results closest to those given above were from the audiovisual mode, whereas by contrast the video-only mode showed relatively less use of the ethnicity-nonstandardness scales.

In sum, the results indicated a positive answer to the first question—the two judgmental dimensions of *ethnicity-nonstandardness* and *confidence-eagerness* were found relevant to the situation where teacher-raters saw the children as well as heard their speech.

Relations of Judgments to Stereotypes

In subsequent research, the second question was approached by having teacher-raters judge children who were not seen or heard but described for them in a stereotyped fashion. Ratings of the children as imagined from these descriptions were then compared with ratings of actual speech sam-

[8] The value in parenthesis is a correlation coefficient. These have a range from 0.0, or no correlation, to 1.0, or perfect correlation. Accordingly, the values reported here are markedly high.

ples (videotapes). Six different stereotype descriptions were prepared, one for each category of children in the videotape samples. As examples, the descriptions for a low-status Anglo and a middle-status Black child were as follows:

> He is Anglo and lives with his mother who is a laundress and his three brothers and one sister in a lower class neighborhood.
>
> He is Black and the son of a professor of sociology at the University. He lives in an upper-middle-class neighborhood.

Ratings of all six stereotype descriptions were obtained from each teacher-rater approximately one week prior to, and one week after ratings of actual speech samples. Again ratings were also obtained for the videotapes of the six categories of children, but this time results were analyzed directly in terms of the scales constituting the two factor model.

Results indicated a definite, but only moderate, statistical relation between ratings of the stereotypes and ratings of the children. That is to say, as a teacher tended, for example, to rate the Anglo samples as relatively high on the confidence-eagerness scales or low in ethnicity-nonstandardness, she *tended* to rate the actual (videotape) sample accordingly.

The implication of the foregoing relation and tentative answer to the second question, was that rather than reporting only a stereotype when rating a child, a teacher-rater may have been rating the videotape samples *relative to* her stereotypes. That is to say, stereotypes may mediate in the differentiation of the speech samples, but the teacher-rater nevertheless is still somewhat sensitive to individual differences *within* presumed categories of children grouped on an ethnicity-by-status basis.

Although the following is properly the subject of further research, it may be that a teacher's sensitivity to individual differences among children in the above categories is reflected in her tendency to exercise greater differentiation in rating actual speech samples relative to the ratings given for stereotypes. Conversely, the less sensitive to individual differences a teacher is to children of a particular category the more the ratings of actual children may be undifferentiated from a stereotype. Put in anecdotal terms, this latter example constitutes a case of the "they all look (or sound) alike" attitude.

Implications

Although studies in the current program of research continue, the findings to date suggest a number of implications. Some of these are, of course, theoretical, but for this report the practical will be emphasized.

One significant shift of thought challenging English language instruction in the schools today is that *differences* in language habits, particularly those of minority group children, have been too often confused as deficits.[9] The fact that, for example, teachers in this research program have consistently based about half of their judgmental perspective upon *nonstandardness* is symptomatic of a prescriptionist (for standard English) rather than, say, an aptness or a communicativeness criterion in evaluating children's speech. This hints of a major shortcoming in what teachers are taught about the language of school children, one of where the ends in teaching English overshadow the means. Perhaps too much is stressed about the objective of teaching (and expecting) standard English rather than the

[9] This issue is discussed further in Frederick Williams, "Language, Attitude, and Social Change," in **Language and Poverty: Perspectives on a Theme**, ed. Frederick Williams (Chicago, 1970).

careful diagnosis of existing linguistic capabilities of children as a starting point. The designation of *nonstandard* (or particularly as some say, *substandard*) implies a classification of "deficiency" in a child's speech which overlooks that a child speaking a nonstandard dialect of English may be as developed, psycholinguistically at least, as his standard-English-speaking age mate. To emphasize the point, this is *not* to argue against the merits of standard English as an instructional objective in American schools, but to stress that teachers might benefit from knowing more about language differences in children as a means for improving English instruction. If only for defining an instructional starting point, an ability to diagnose what a child can do linguistically in nonstandard English should introduce some efficiencies into English language instruction.

Just as the present studies imply a bias in teachers' attitudes toward nonstandardness, they also suggest ways to measure such bias and still more to gauge the effects of teacher training.

As mentioned earlier, teachers' stereotype ratings appear quite consistent and do seem to influence judgments of actual speech. If these stereotypes somehow serve as judgmental "anchors," then effectiveness of teacher training in language differences might be gauged by shifts of the anchor point as well as by increased differentiation of actual speech ratings about that point. In all, the instructional goal in teacher training would be one of sensitivity to, and objective differentiation of, language differences. The goal is to reduce the effects of a teacher's stereotype bias in evaluating the language of her pupils.

As Eliza Doolittle counseled Professor Higgins, it is how you treat people that makes them what they are to you. The same advice seems pertinent to the reduction of teachers' negative stereotypes of children who speak nonstandard English.

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED
BY AUTHOR AND THE "ENGLISH RECORD"
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."